Applying Machine Learning to MOMA Science Data for Science Autonomy

Victoria Da Poian[1,2] (victoria.dapoian@nasa.gov), Eric I. Lyness[1,3] (eric.i.lyness@nasa.gov),

Ryan M. Danell[4] , Melissa G. Trainer[1] , Xiang Li[1,5], William B. Brinckerhoff[1] and the MOMA team

[1]Planetary Environments Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD; [2]SURA CRESST, DC; [3]Microtel LLC, MD; [4]Danell Consulting, Inc., Winterville, NC; [5]University of Maryland, Baltimore County, MD

The Mars Organic Molecule Analyser Instrument

Poster ID: 54

Space Exploration Context

Ground-in-the-loop limitations
Remote destinations and shorter at-target mission lifetimes limit or preclude ground-in-the-loop interactions

Communication challenges
Remote destinations and extreme environments involve longer communication delays and smaller data downlink capacities

Detection challenges
Scientists will not be able to guide spacecrafts' instrumentation in detection opportunistic features of interest

Data prioritization
Future instruments will certainly generate more data: data prioritization is vital to optimize mission science return

Science Autonomy

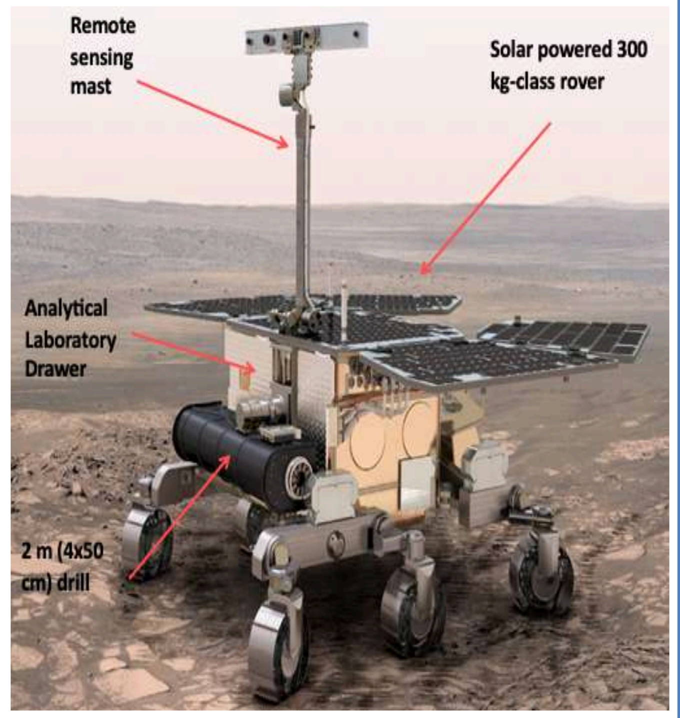
The ability of a science instrument to **analyze its own data** in order:
- to **calibrate** itself
- **optimize ops parameters** based on real-time findings
- **make mission-level decisions** based on scientific observations
- determine which **data** products to **prioritize** and send back first

ExoMars Rover Mission

Objectives

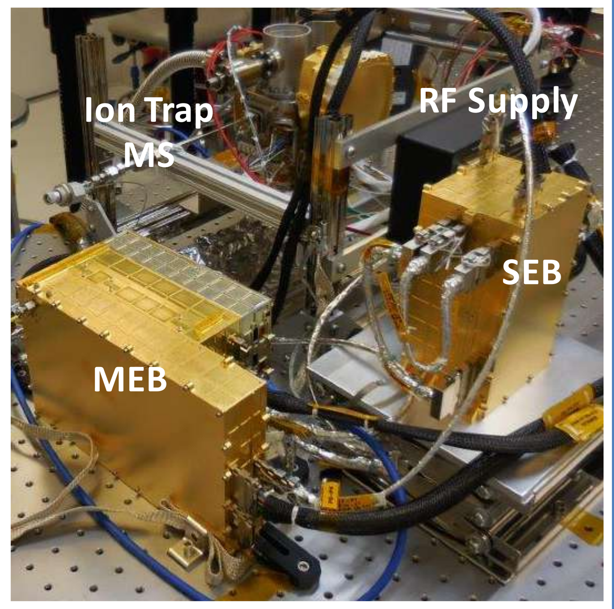
- To search for signs of past and present life on Mars
- To investigate the water/geochemical environment as a function of depth in the shallow subsurface

Instruments



- Drill delivers samples from 2m below the surface to the crushing station
- **Mars Organic Molecule Analyzer (MOMA):** Linear Ion Trap Mass Spectrometer
- **Raman Laser Spectrometer (RLS):** spectral analysis
- **MicrOmega:** imaging the samples (near IR hyperspectral microscope), mineral identification informs MOMA and RLS of regions of interest to target

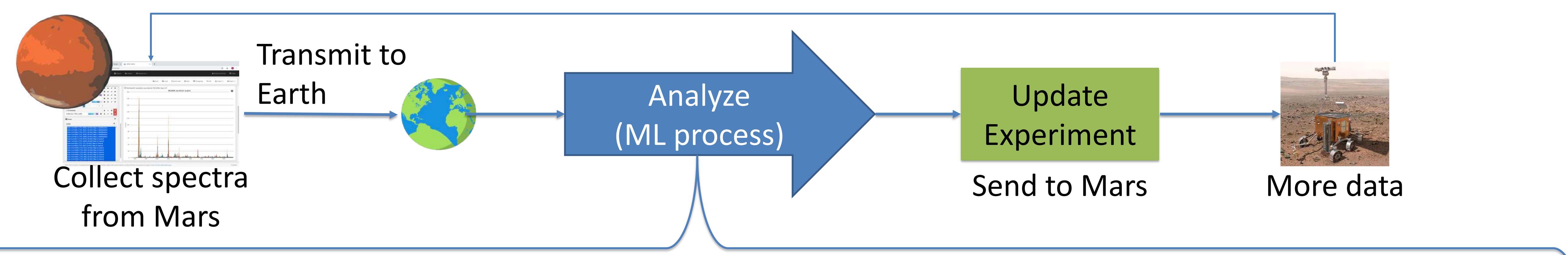
MOMA instrument

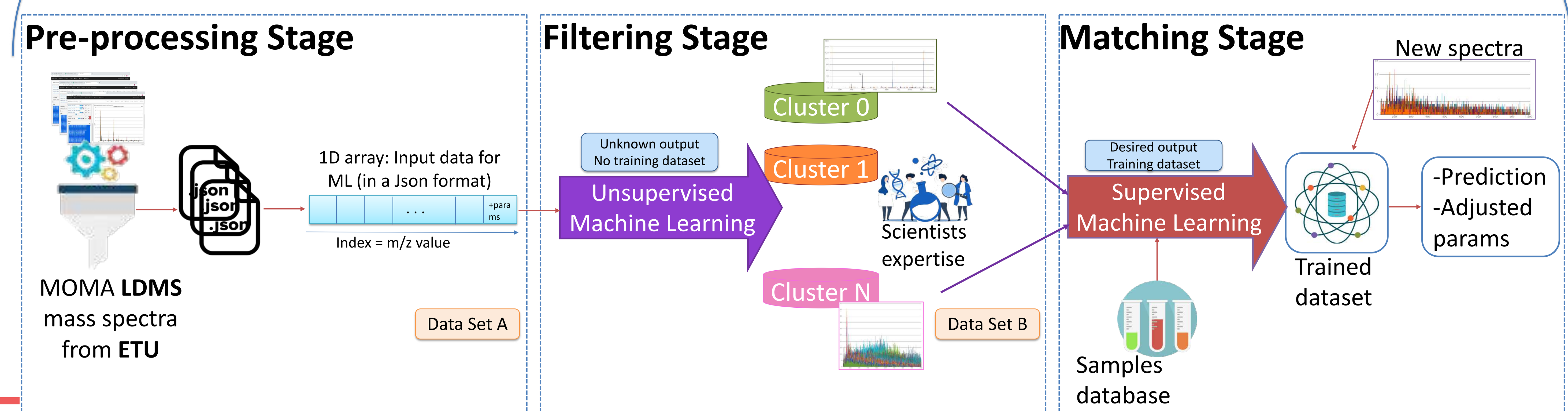


- Dual-source linear ion trap mass spectrometer coupled to pyrolysis/ derivatization-GC (GCMS mode) and UV laser for desorption / ionization (LDMS mode) of crushed rock samples
- Seeks the *molecular* signs of life with *broad* sensitivity to organics and analysis of chirality

Machine Learning process

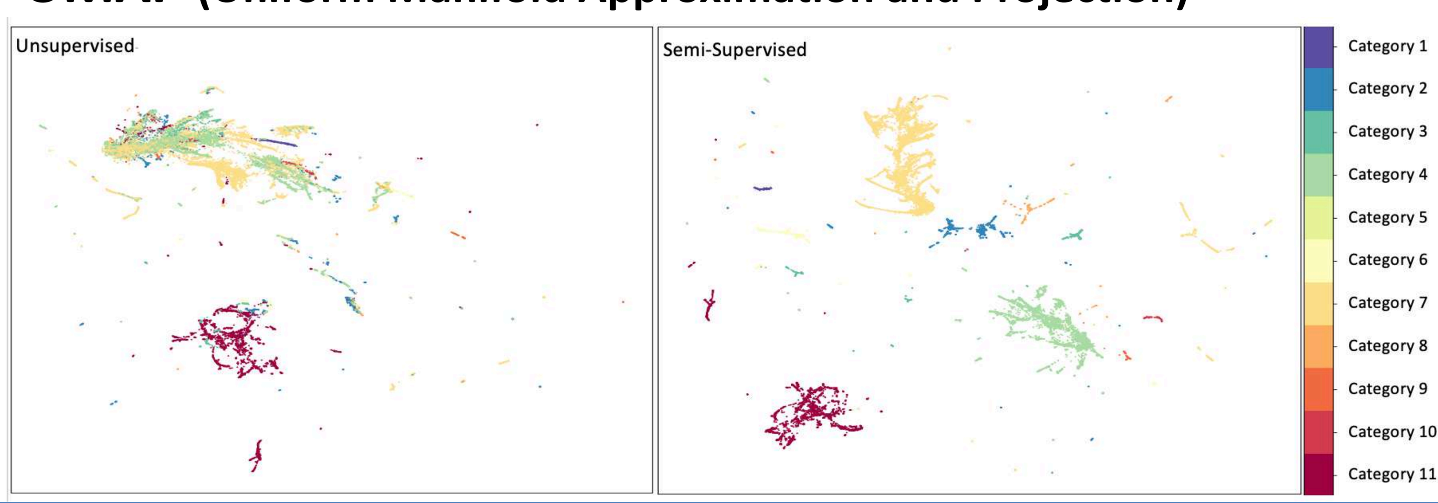
Motivation: The MOMA science team may only have a few hours to analyze the delivered data from Mars and to determine what further experiments should be done to meet the mission's science goals. We are investigating the use of ML to help the science team by matching Flight Model (FM) data from Mars to similar data from tests performed with the Engineering Test Unit (ETU) on Earth.





Initial results

UMAP (Uniform Manifold Approximation and Projection)



- Assists in **high dimension data visualization**
- Provides a **Semi-Supervised** method to help further cluster our data
- Possibility to add **new unseen data** into an existing embedding space
- UMAP Reduced data can be used as a **pre-processing** step of Supervised Learning

<u>**Other results:**</u> (feel free to contact me)

- **Data Processing:** dimensionality reduction, outlier detection,
- **Filtering Stage:** clustering algorithms
- **Matching Stage:** neural network development, implementation and CAL interface

Key Lessons

Data Volume
Challenging acquisition of large datasets acquisition for ML trainings

Team Efforts
Crucial collaboration between scientists and data science team

Resource Limits
Desired performance and available resources tradeoffs (CPU, memory)

Trusting ML
Dev. of a "Trust Readiness Level" index (like TRL in engineering dev.)